Proceeding Paper

Source Camera Linking Algorithm Based on the Analysis of Plain Image Zones †

Ana Elena Ramirez-Rodriguez *, Mariko Nakano and Hector Perez-Meana

Instituto Politécnico Nacional, Av. Luis Enrique Erro S/N, Unidad Profesional Adolfo López Mateos, Zacatenco, Alcaldía Gustavo A. Madero, Ciudad de México 07738, Mexico; mnakano@ipn.mx (M.N.); hmperezm@ipn.mx (H.P.-M.)

* Correspondence: aramirezr0906@alumno.ipn.mx
† Presented at the 4th International Conference on Communications, Information, Electronic and Energy Systems (CIEES 2023), Plovdiv, Bulgaria, 23–25 November 2023.

Abstract: This paper proposes an image linking scheme, where the plain zones of the image under analysis are used, which are detected using the Discrete Cosine Transform. Then, a Block Matching 3D denoising filter is used to reduce the additive noise such that the residual error of the plain zones. The Peak to Correlation Energy is applied to compare the enhanced residual noise patterns. Experimental results show that the proposed methodology compared to conventional techniques is better due to the emphasis on plain zones, because the detailed and edge areas introduce distortions that can significantly affect the extracted residual noise patterns.

Keywords: digital forensic; camera identification; denoising filter; camera linking; plain detection

1. Introduction

In recent years, the development of efficient software for image tampering and editing has made image manipulation more accessible, even for non-specialists. This technological progress has created new possibilities for creative expression and image enhancement. As a consequence, camera identification and the verification of the authenticity of images are more difficult. For this reason, image forensics has become an important research field. Also, some forensic techniques are aimed at the detection of illegal activities. These methods ensure image integrity and identify the source of digital images.

One of the main tasks in image forensics is source image identification (SCI) [1–4]. We can consider two different problems. In the first one, it consists of determining if the image was taken using a specific camera. The second one is Source Camera Linking (SCL) which is focused on determining if both images are from the same camera or not. Among the several features that can be extracted from a digital image to solve this problem, the Photo Response Non-Uniformity (PRNU), which is a fingerprint of the camera, appears to be a suitable approach due to the imperfection introduced to the sensor during fabrication and the inhomogeneity of the sensor wafers [1], which can be considered a deterministic Sensor Pattern Noise (SPN).

Several methods have been proposed to estimate the PRNU or the SPN, which requires the use of a denoising filter [1–3], among them the Mihcak filter [5,6] and BM3D [7,8], although a comparison between both methods with 18 different cameras showed that the BM3D filter provides a better performance [7,8].

Usually, the SCL problem is solved using only the residual noise extracted for each image. The most common method is cropping a zone in the centre of the image, creating a smaller region for its analysis that contains detailed and texture zones. Thus, the pattern noise causes some errors since the SCL schemes produce false-positive and false-negative results. These inaccuracies can lead to incorrect identification. For this reason, in this paper, we propose a SCL algorithm that first analyses the image content, identifying the plain

regions through a block classification algorithm [9]. Next, the BM3D denoising filter is applied to extract the pattern noise fingerprint. Then, the cross-correlation is calculated using the PCE. The evaluation results, which focus on mainly plain areas, demonstrate that the proposed scheme provides better results than conventional methods.

The rest of the paper is organised as follows: Section 2 presents a description of the proposed method. In Section 3, the experimental results are shown, and finally, in Section 4, the conclusions of this research.

2. Proposed Method

Figure 1 shows the proposed source camera linking method in which firstly, the images $I1(x,y)$ and $I2(x,y)$ are analysed to detect the plain regions presented in both images. The pixels of the plain regions are used for the construction of two masks $Z1(x,y)$ and $Z2(x,y)$ encoding with 1 the position of the pixels belonging to the plain regions of each image and with zero the remaining ones. Next, both masks are multiplied among them to build a mask $Zc(x,y) = Z1(x,y) \, Z2(x,y)$ which is used to determine the position of common plain regions in both images. Then, the common plain regions in both images are obtained by multiplying each image by $Zc(x,y)$. The estimated plain regions are fed into a denoising filter, whose output is used to estimate the residual noise. The residual noises are then used to estimate the PCE, which is used to determine if both images belong to the same camera.
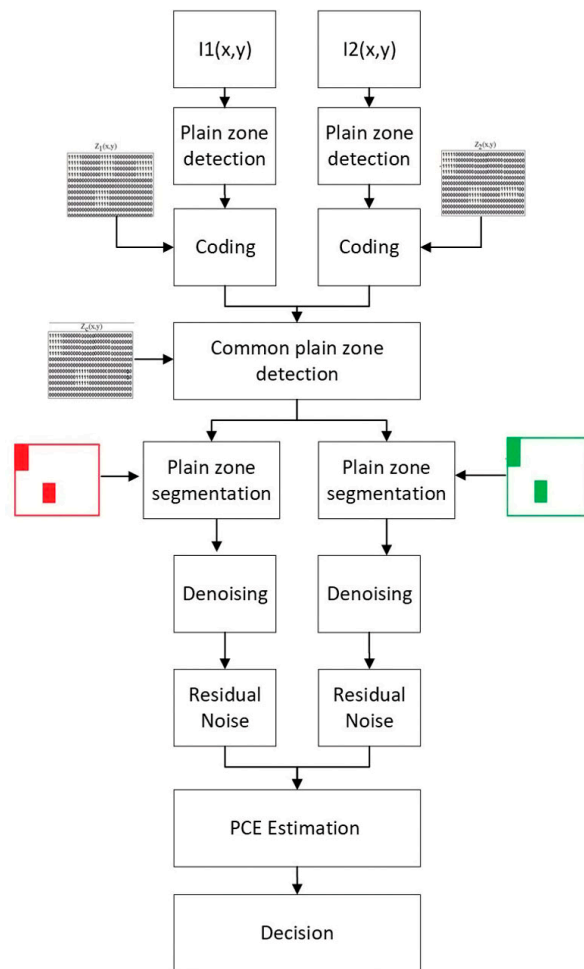

Figure 1. Proposed source camera linking method.

2.1. Plain Zone Segmentation

Obtaining a clean characteristic noise as a pattern may not be easy because, in practice, this desired noise can be contaminated by some distorted factors, such as the image content,

including edges and textures, which lead to distortions in the noise patterns. Consequently, it becomes crucial to identify the plain regions in the image that are minimally affected by such distortions. Therefore, it is fundamental to determine the image region, which may contain lower distortions that may result in a contaminated residual noise pattern.

To avoid this problem, the proposed system estimates the residual noise using only the plain zones on the images under analysis. They are estimated using the zone classification scheme shown in Figure 2, where the image zones are classified as plain, edge, and texture.
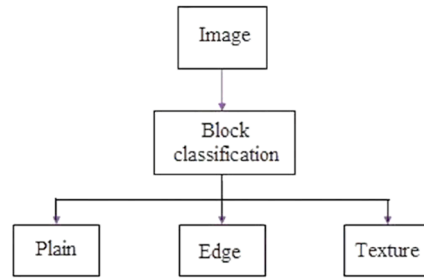

Figure 2. Zone classification scheme.

The plain zones are estimated using the classification scheme presented in [10]. The block classification is obtained using the Discrete Cosine Transformation (DCT), which segments the image into non-overlapped blocks of 8 × 8 pixels, whose coefficients are classified as shown in Figure 3. Using the frequencies estimated in each block, the following variables are determined.

Figure 3. Block classification diagram.

Besides *L*, *H* and *E* given by (1) and (3), several other conditions are proposed to determine the block type [10] as follows: $\mu1 = 125$, $\mu2 = 900$, $\alpha1 = 2.3$, $\alpha2 = 1.4$, $\beta1 = 1.6$, $\beta2 = 1.1$, $\gamma = 4$ and $\kappa = 290$.

$$L = \sum_{k=1}^{2} I_{dc}(0,k) + \sum_{k=0}^{2} I_{dc}(0,k) + I_{dc}(2,0); \tag{1}$$

$$E = \sum_{k=3}^{6} I_{dc}(k,0) + \sum_{k=3}^{6} I_{dc}(0,k) + \sum_{k=1}^{2} I_{dc}(3,k) + I_{dc}(3,3); \tag{2}$$

$$H = \sum_{k=0}^{7} \sum_{m=0}^{7} I_{dc}(k,m) - E - L. \tag{3}$$

Using these conditions together with the values given by Equations (1)–(3) the block characteristics are determined using the magnitude of the ratios $(L + E)/H$ and (L/E) which indicate the presence of an edge while $E + H$ is an approximation for a texture block. Thus, a plain block is detected if $E + H \leq \mu1$, while an edge block is detected if any of the next conditions are satisfied:

- if $E + H \leq \mu 2$ and

$$
\begin{cases}
\frac{L}{E} \geq \propto 1 & \text{and} \quad \frac{L+E}{H} \geq \beta 1 \\
\frac{L}{E} \geq \beta 1 & \text{and} \quad \frac{L+E}{H} \geq \alpha 1 \quad ; \\
& \frac{L+E}{H} \geq \gamma
\end{cases}
\tag{4}
$$

- or if $E + H > \mu 2$ and

$$
\begin{cases}
\frac{L}{E} \geq \propto 2 & \text{and} \quad \frac{L+E}{H} \geq \beta 2 \\
\frac{L}{E} \geq \beta 2 & \text{and} \quad \frac{L+E}{H} \geq \alpha 2 \quad . \\
& \frac{L+E}{H} \geq \gamma
\end{cases}
\tag{5}
$$

Finally, a texture block is detected if the block does not satisfy any of the above conditions and $E + H > \kappa$. Figure 4 shows the performance of the proposed scheme, where Figure 4a shows an example of an original image and Figure 4b shows the classification result using the above-described method, where the white colour represents the plain blocks, the green colour denotes the edges and the blue is the texture blocks.
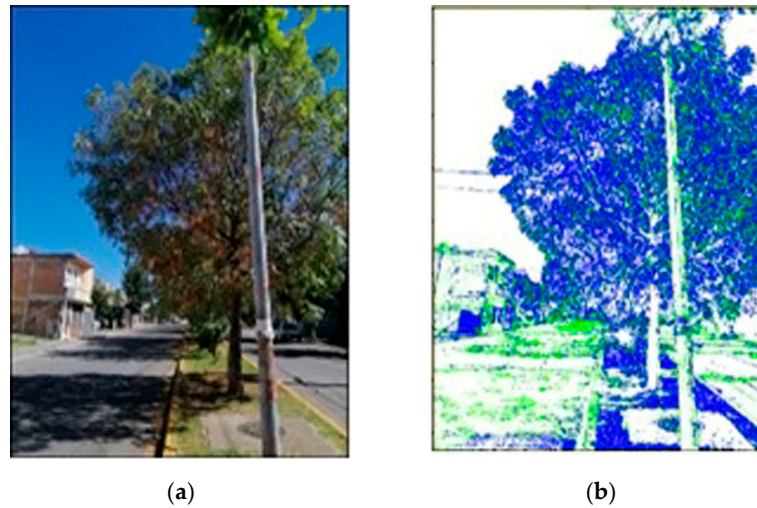

(**a**)   (**b**)

Figure 4. Example of an image: (**a**) original image; (**b**) after applying block classification.

Once the block classification is achieved, it is necessary to determine which image region is more suitable for estimating the residual noise. As an example, we applied the block detection method described above to Figure 5, which contained plain zones, edges, and textures, obtaining the block classification shown in Figure 6a, where the white colour denotes the plain blocks, the green colour represents the edges, and the blue colour represents the texture blocks, respectively. Next, the residual noise shown in Figure 6b was estimated. The noise pattern extracted from the plain blocks does not present a visible distortion, while the noise pattern extracted from the edge and texture blocks is highly distorted. This fact suggests that the noise pattern extracted from plain blocks provides a more reliable estimation of residual noise pattern than those extracted from zones including edge and texture blocks, as shown in Figure 7.
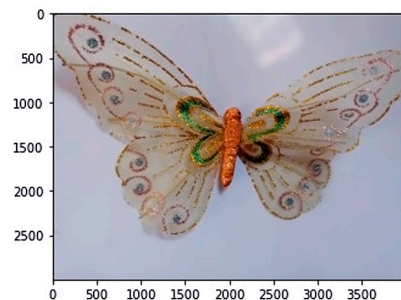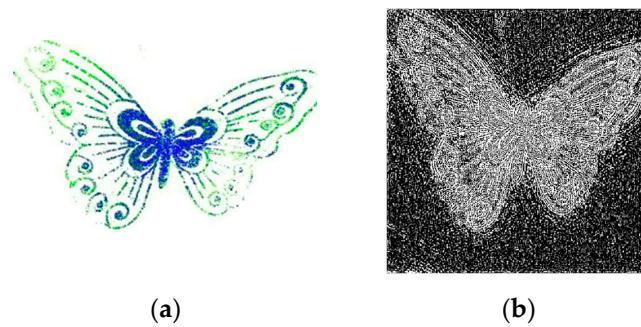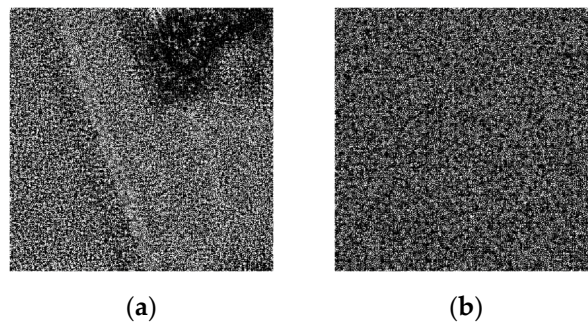

Figure 5. Example of the original image.

Figure 6. Image obtained after using the proposed scheme: (**a**) block classification; (**b**) residual noise estimated from plain, edge and texture zones.

Figure 7. Residual noise obtained from the image under analysis: (**a**) the region contains textures and edges; (**b**) the region is plain.

To improve the performance of the proposed scheme, the residual noise will be extracted only from plain blocks instead of a zone located in the centre of the image, as usually done. Thus, to carry out the image linking, once the plain areas of images are under analysis, I_i, $i = 1$, 2, are identified, and the pixels belonging to them are encoded, as in (6):

$$z_i(x,y) = \begin{cases} 1, \text{if } (x,y) \text{ is plain} \\ 0, \text{Otherwise} \end{cases}. \tag{6}$$

Subsequently, the common plain blocks are segmented in both images, described in Equation (7):

$$Ic_i(x,y) = z_1(x,y)z_2(x,y)I_i(x,y); i = 1, 2. \tag{7}$$

Thus, the plain blocks of size N × N obtained using (7) are used to feed into the denoising stage. If two or more regions in both images have a suitable size to carry out the camera linking, the first section that was segmented is selected; otherwise, several plain regions can be concatenated to have a plain region of a suitable size.

2.2. Estimation of Peak to Correlation Energy

Once the plain zones with the same size and location in both images are detected, they are fed into the image denoising stage, whose outputs are used to improve the residual noise estimation [5,7]. Several denoising filters have been proposed; among them, a suitable approach to reduce the noise in the image is the BM3D filter [7], which consists of two stages: The first stage estimates the denoised image using a thresholding, which is part of the collaborative filter. The second step uses both the original noisy image and the image obtained in the first step, together with a Wiener filter [8]. Because several evaluation results have shown that the BM3D provides better performance [8,11], when used in the denoising stage of several image processing applications, it will be used to estimate the image noise pattern in the decision stage as follows:

$$n_i = Ic_i - Is_i, \tag{8}$$

where Ic_i is the segmented plain block of the *i-th* image and Is_i is the enhanced plain block of the same image obtained using the BM3D denoising filter. Next, after the residual noises are estimated in (9), they are enhanced, as proposed in [12]:

$$n_{ei}(x,y) = \begin{cases} e^{-0.5n_i^2(x,y)/\alpha^2}, \text{if } 0 \leq n_i(x,y) \\ -e^{-0.5n_i^2(x,y)/\alpha^2}, \text{otherwise} \end{cases}. \tag{9}$$

The next step is to obtain the PCE, which compares the residual noise patterns extracted after the noise enhancement stage, which is given by (10):

$$PCE = \frac{NCC^2(0,0)}{\frac{1}{MN}\sum_i \sum_j NCC^2(i,j)}, \tag{10}$$

where

$$NCC(i,j) = \frac{\sum_{k=0}^{N-1} \sum_{i=0}^{M-1} \left[ne_1(i,j) - \overline{ne}_1\right]\left[ne_2(i,j) - \overline{ne}_2\right]}{\left\|ne_1(i,j) - \overline{ne}_1\right\|\left\|ne_2(i,j) - \overline{ne}_2\right\|}. \tag{11}$$

The PCE calculates the height of the peaks and the energy obtained from the cross-correlation between the patterns of the two characteristic noises. Therefore, if the value of the PCE is large, there is a greater possibility that the images belong to the same camera.

3. Results

To test the proposed method, we used the *Forchheim Image Database (FODB)* [13]. This database contains an extensive collection of 3851 images from the native camera, with outdoor and indoor, day and night scenarios, also including horizontal and vertical image orientations; all cameras were setting to automatic mode. We focused on testing our implementation only using the horizontal images from 12 cameras. The tested cameras are presented in Table 1; an example of a tested image is shown in Figure 8.

Table 1. Database information.

Number of Cameras	Brand	Model	OS	Resolution
Camera 1	Motorola	E3	Android	3280 × 2664
Camera 2	LG	Optimus L50	Android	2048 × 1536
Camera 3	Wiko	Lenny 2	Android	2560 × 1920
Camera 4	LG	G3	Android	4160 × 3120
Camera 5	Apple	iPhone 6s	iOS	4032 × 3024
Camera 6	LG	G6	Android	2080 × 1560
Camera 7	Motorola	Z2 Play	Android	4032 × 3024
Camera 8	Motorola	G8 Plus	Android	3000 × 4000
Camera 9	Samsung	Galaxy S4 mini	Android	3264 × 2448
Camera 10	Samsung	Galaxy J1	Android	2592 × 1944
Camera 11	Samsung	Galaxy J3	Android	3264 × 2448
Camera 12	Samsung	Galaxy Star 5280	Android	1600 × 1200

Figure 8. Example of a tested image.

Also, the experiments were carried out to test the proposed algorithm. The hardware characteristics were NVIDIA GeForce GTX 1650, Intel® Core™ i7 10th gen; additionally, the proposed algorithm was implemented in Python 3.8.5.

3.1. Experiment I

In experiment I, the PCE was obtained when all tested images were from the same camera source. The images were cropped in two different ways:

- The first way the image was cropped was by a single 512×512 plain pixel region, named 'one zone';
- The second way consisted of cropping sixteen plain sections of 128×128 pixel size, which were reassembled into a 512×512 pixel image, referred to as 'zones'.

Both methods were compared against the conventional method, where the cropping of the 512×512 pixel zone was in the centre. Table 2 shows the PCE obtained.

Table 2. PCE results from the same camera.

Number of Cameras	One Zone	Zones	Conventional
Camera 1	35.1152	42.1491	6.4015
Camera 2	37.1227	44.2302	9.0886
Camera 3	33.6813	42.7957	7.6214
Camera 4	34.5184	45.0543	5.2810
Camera 5	35.4917	45.0038	6.9693
Camera 6	33.7016	40.7761	4.7675
Camera 7	36.0282	49.5316	8.7290
Camera 8	35.5511	44.6075	3.8505
Camera 9	34.3121	43.5091	5.6535
Camera 10	33.4316	41.5974	4.0072
Camera 11	35.9563	49.4259	7.5803
Camera 12	35.5768	43.1329	3.5796

The results of Experiment I showed that the PCE obtained with the proposed method was higher than in the conventional method, so the SCL would be more accurate due to the fact that the larger the PCE value is, the more likely it is that the images belong to the same camera.

3.2. Experiment II

The second experiment consisted of obtaining the PCE when the images from one camera were compared against the images from the other cameras as we can see in Table 3, again by performing the two types of cropping, as mentioned in Experiment I.

Table 3. PCE results from different camera.

Number of Cameras	One Zone	Zones	Conventional
Camera 1	0.7037	0.6353	0.9974
Camera 2	0.5950	0.5837	0.8295
Camera 3	0.6174	0.5993	0.9791
Camera 4	0.4690	0.3725	0.6912
Camera 5	0.5424	0.5278	0.7859
Camera 6	0.4785	0.3942	0.9171
Camera 7	0.6315	0.6138	0.8450
Camera 8	0.5101	0.4496	0.7488
Camera 9	0.5673	0.4569	0.8922
Camera 10	0.5415	0.4110	0.9822
Camera 11	0.5259	0.4912	0.7551
Camera 12	0.4772	0.2951	0.8825

On the other hand, Experiment II shows that the PCE results obtained with the proposed method were smaller than those obtained with the conventional method, as the images were taken from different cameras; i.e., the pattern noise was different and unique for each camera, so the closer to zero the PCE result is, the more accurate it will be to determine that the images do not belong to the same camera.

4. Conclusions

In this paper, we proposed a method to improve the source camera linking by analysing only the plain areas of the images, as we realised how the noise caused by the detail and edge areas distorts the pattern noise affecting the correct camera source linking, while the conventional method does not pay attention to the image features such as edges and textures, so this can cause false positive or false negative results at the moment to determine if the images belong to the same camera or not. As can be seen, the proposed method is better to determine whether or not the images belong to the same camera. The use of plain zones reduces noise distortions to increase the robustness and efficiency of the proposed system. It is necessary to highlight that our method takes advantage of the detection of plain zones and noise enhancement since conventional algorithms focus on regions that contain edges or detailed zones, which distort the extracted noise pattern.

Author Contributions: Conceptualization, A.E.R.-R., M.N. and H.P.-M.; methodology, A.E.R.-R., M.N. and H.P.-M.; software, A.E.R.-R., M.N. and H.P.-M.; validation, A.E.R.-R., M.N. and H.P.-M.; formal analysis, A.E.R.-R., M.N. and H.P.-M.; investigation, A.E.R.-R., M.N. and H.P.-M.; resources, A.E.R.-R., M.N. and H.P.-M.; data curation, A.E.R.-R., M.N. and H.P.-M.; writing—original draft preparation, A.E.R.-R., M.N. and H.P.-M.; writing—review and editing, A.E.R.-R., M.N. and H.P.-M.; visualization, A.E.R.-R., M.N. and H.P.-M. All authors have read and agreed to the published version of the manuscript.

References

1. Lukás, J.; Fridrich, J.; Goljan, M. Digital Camera Identification from Sensor Pattern Noise. *IEEE Trans. Inf. Forensics Secur.* **2006**, *1*, 205–214. [CrossRef]
2. Chen, J.M.; Fridrich, J.; Goljan, M.; Jan, L. Determining Image Origin and Integrity Using Sensor. *IEEE Trans. Inf. Forensics Secur.* **2008**, *3*, 74–90. [CrossRef]
3. Hu, Y.; Yu, B.; Jian, C. Source Camera Identification Using Large Components of Sensor Pattern Noise. In Proceedings of the 2nd International Conference on Computer Science and its Applications, Seoul, Republic of Korea, 10–12 December 2009.
4. Tiwari, M.; Gupta, B. Image Features Dependant Correlation-Weighting Function for Efficient PRNU Based Source Camera Identification. *Forensic Sci. Int.* **2018**, *285*, 111–120. [CrossRef] [PubMed]
5. Valesia, D.; Coluccia, G.; Bianchi, T.; Magli, E. User Authentication via PRNU-Based and Physical Unclonable Functions. *IEEE Trans. Inf. Forensics Secur.* **2008**, *12*, 1941–1956. [CrossRef]
6. Mihcak, M.k.; Kozintsev, I.; Ramchandran, K. Spatially Adaptive Statistical Modeling of Wavelet Image Coefficients and Its Application to Denoising. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ, USA, 15–19 March 1999.
7. Mihcak, M.k.; Kozintsev, I.; Ramchandran, K.; Moulin, P. Low-Complexity Image Denoising Based on Statistical Modeling of Wavelet Coefficients. *IEEE Signal Process. Lett.* **1999**, *6*, 300–303. [CrossRef]

8. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Trans. Inf. Forensics Secur.* **2007**, *1*, 2080–2095. [CrossRef] [PubMed]

9. Salazar, D.A.; Ramirez-Rodriguez, A.E.; Nakano, M.; Cedillo-Hernandez, M.; Perez-Meana, H. Evaluation of Denoising Algorithms for Source Camera Linking. In Proceedings of the MCPR 2021, Mexico City, Mexico, 19–22 June 2021.

10. Tong, H.H.Y.; Venetsanopoulos, A.N. A perceptual model for JPEG applications based on block classification, texture masking, and luminance masking. In Proceedings of the International Conference on Image Processing, Chicago, IL, USA, 7 October 1998.

11. Su, Q.; Wang, Y.; Li, Y.; Zhang, C.; Lang, P.; Fu, X. Image Denoising Based on Wavelet Transform and BM3D Algorithm. In Proceedings of the IEEE 4th International Conference on Signal and Image Processing, Wuxi, China, 19–21 June 2019.

12. Li, C.T. Source Camera Identification Using Enhanced Sensor Pattern Noise. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 280–287.

13. Hadwiger, B.; Riess, C. The Forchheim Image Database for Camera Identification in the Wild. In Proceedings of the ICPR Workshops, Milano, Italy, 10–15 January 2021.